

Protein folding and binding can emerge as evolutionary spandrels through structural coupling

Michael Manhart¹ and Alexandre V. Morozov^{1,2*}

¹*Department of Physics and Astronomy and* ²*BioMaPS Institute for Quantitative Biology,*
Rutgers University, Piscataway, NJ 08854, USA

Abstract

Binding interactions between proteins and other molecules mediate numerous cellular processes, including metabolism, signaling, and regulation of gene expression. These interactions evolve in response to changes in the protein’s chemical or physical environment (such as the addition of an antibiotic), or when genes duplicate and diverge. Several recent studies have shown the importance of folding stability in constraining protein evolution. Here we investigate how structural coupling between protein folding and binding – the fact that most proteins can only bind their targets when folded – gives rise to evolutionary coupling between the traits of folding stability and binding strength. Using biophysical and evolutionary modeling, we show how these protein traits can emerge as evolutionary “spandrels” even if they do not confer an intrinsic fitness advantage. In particular, proteins can evolve strong binding interactions that have no functional role but merely serve to stabilize the protein if misfolding is deleterious. Furthermore, such proteins may have divergent fates, evolving to bind or not bind their targets depending on random mutation events. These observations may explain the abundance of apparently nonfunctional interactions among proteins observed in high-throughput assays. In contrast, for proteins with both functional binding and deleterious misfolding, evolution may be highly predictable at the level of biophysical traits: adaptive paths are tightly constrained to first gain extra folding stability and then partially lose it as the new binding function is developed. These findings have important consequences for our understanding of fundamental evolutionary principles of both natural and engineered proteins.

Proteins carry out a diverse array of chemical and mechanical functions in the cell, ranging from metabolism to signaling [1]. Thus proteins serve as central targets for natural selection in wild populations, as well as a key toolbox for bioengineering novel molecules with medical and industrial applications [2, 3]. Most proteins must fold into their native state, a unique three-dimensional conformation, in order to perform their function, which

*Corresponding author: morozov@physics.rutgers.edu

typically involves binding a target molecule such as DNA, RNA, another protein, or a small ligand [1]. Misfolded proteins may also form toxic aggregates and divert valuable protein synthesis and quality control resources [4–7]. It is therefore imperative that the folded state be stable against the thermal fluctuations present at physiological temperatures. However, biophysical experiments and computational studies reveal that most random mutations in proteins destabilize the folded state [8, 9], including mutations that improve function [9–11]. As a result many natural proteins tend to be only marginally stable, mutationally teetering at the brink of substantial unfolding [12, 13]. With proteins in such a precarious evolutionary position, how can they evolve new functions while maintaining sufficient folding stability?

Directed evolution experiments have offered a window into the dynamics of this process [2, 3], indicating the importance of compensatory mutations, limited epistasis, and mutational robustness. Theoretical efforts to describe protein evolution in biophysical terms have focused on evolvability [14], global properties of protein interaction networks [15, 16], and reproducing observed distributions of protein stabilities and evolutionary rates [13, 17–19]. However, a subtle but key property of proteins has not been explored in this context: structural coupling of folding and binding (the fact that folding is required for function) implies evolutionary coupling of folding stability and binding strength. Thus selection acting directly on only one of these traits may produce apparent, indirect selection for the other. The importance of this effect was popularized by Gould and Lewontin in their influential paper on evolutionary “spandrels” [20], defined as traits that evolve as byproducts in the absence of direct selection. Since then the importance of coupling between traits has been explored in many areas of evolutionary biology [21], including various molecular examples [12, 22, 23].

How do coupled traits affect protein evolution? We consider a simple model that describes evolution of a new binding interaction in the context of a directed evolution experiment [3], as a result of gene duplication and divergence [24], or in response to a change in the protein’s chemical or physical environment, including availability and concentrations of various ligands [25, 26] as well as temperature [27, 28]. We postulate a fitness landscape as a function of two biophysical traits: stability and the free energy of binding a target molecule. We then use an exact numerical algorithm [29, 30] to quantitatively characterize adaptation on this fitness landscape, addressing key evolutionary questions of epistasis [31, 32], predictability [25, 33, 34], and the tempo of adaptation [17, 35].

Results

Model of protein energetics. We consider a protein with two-state folding kinetics [1]. In the folded state, the protein has an interface that binds a target molecule. Because the protein can bind *only* when it is folded, the binding and folding processes are structurally coupled. Under the thermodynamic equilibrium assumption (valid when protein folding and binding are faster than typical cellular processes), the probabilities of the three structural states – folded and bound ($p_{f,b}$), folded and unbound ($p_{f,ub}$), and unfolded and unbound ($p_{uf,ub}$) – are given by their Boltzmann weights:

State	Free energy	Probability
folded, bound	$E_f + E_b$	$p_{\text{f,b}} = Z^{-1}e^{-\beta(E_f+E_b)}$
folded, unbound	E_f	$p_{\text{f,ub}} = Z^{-1}e^{-\beta E_f}$
unfolded, unbound	0	$p_{\text{uf,ub}} = Z^{-1}$

(1)

Here β is the inverse temperature, E_f is the free energy of folding (also known as ΔG), and $E_b = E'_b - \mu$, where E'_b is the binding free energy and μ is the chemical potential of the target molecule. For simplicity, we will refer to E_b as the binding energy. Note that $E_f < 0$ for intrinsically-stable proteins and $E_b < 0$ for favorable binding interactions. The partition function is $Z = e^{-\beta(E_f+E_b)} + e^{-\beta E_f} + 1$.

The folding and binding energies depend on the protein’s genotype (amino acid sequence) σ . We assume that adaptation only affects “hotspot” residues at the binding interface [36, 37]; the rest of the protein does not change on relevant time scales because it is assumed to be already optimized for folding. If positions away from the binding interface can accept stabilizing mutations (and are not functionally constrained), they may be explicitly included into the model as “folding hotspots.” In the present study we focus on L binding hotspot residues which, to a first approximation, make additive contributions to the total folding and binding free energies [38] (see SI Methods for the discussion of non-additive effects):

$$E_f(\sigma) = E_f^{\text{ref}} + \sum_{i=1}^L \epsilon_f(i, \sigma^i), \quad E_b(\sigma) = E_b^{\text{min}} + \sum_{i=1}^L \epsilon_b(i, \sigma^i), \quad (2)$$

where $\epsilon_f(i, \sigma^i)$ and $\epsilon_b(i, \sigma^i)$ capture the energetic contributions of amino acid σ^i at position i . The reference energy E_f^{ref} is the fixed contribution to the folding energy from all other residues in the protein. Furthermore, by construction it is also the total folding energy of a reference sequence σ_{ref} (see Methods), so that each $\epsilon_f(i, \sigma^i)$ can be interpreted as the change in total folding free energy E_f ($\Delta\Delta G$ value) resulting from a single-point mutation of σ_{ref} . The parameter E_b^{min} is the minimum binding energy among all genotypes (see Methods). Amino acid energies $\epsilon_f(i, \sigma^i)$ and $\epsilon_b(i, \sigma^i)$ are randomly sampled from distributions constructed using available $\Delta\Delta G$ data and other biophysical considerations (see Methods); the exact shape of these distributions is unimportant for large enough L due to the central limit theorem.

Fitness landscape. We construct a simple fitness landscape based on the molecular traits E_f and E_b . Without loss of generality, we assume that the protein contributes fitness 1 to the organism if it is always folded and bound. Let $f_{\text{ub}}, f_{\text{uf}} \in [0, 1]$ be the multiplicative fitness penalties for being unbound and unfolded, respectively: the fitness is f_{ub} if the protein

is unbound but folded, and $f_{\text{ub}}f_{\text{uf}}$ if the protein is both unbound and unfolded. Then the fitness of the protein averaged over all three possible structural states in Eq. 1 is given by

$$\mathcal{F}(E_f, E_b) = p_{\text{f,b}} + f_{\text{ub}}p_{\text{f,ub}} + f_{\text{ub}}f_{\text{uf}}p_{\text{uf,ub}}. \quad (3)$$

This fitness landscape is divided into three nearly-flat plateaus corresponding to the three protein states of Eq. 1, separated by steep thresholds corresponding to the folding and binding transitions (Fig. 1A). The heights of the plateaus are determined by the values of f_{ub} and f_{uf} , leading to three qualitative regimes of the global landscape structure (Fig. 1B–D).

In the first case (Fig. 1B), a protein that is perfectly folded but unbound has no fitness advantage over an unbound and unfolded protein: $f_{\text{ub}} = f_{\text{ub}}f_{\text{uf}}$. Thus selection acts directly only on the binding trait. This regime requires that either $f_{\text{ub}} = 0$ (binding is essential, e.g., in the context of conferring antibiotic resistance to the cell [25]) or $f_{\text{uf}} = 1$ (misfolded proteins are not toxic). The latter case also includes directed evolution experiments where only function is artificially selected for *in vitro*. In contrast, when $f_{\text{ub}} = 1$ and $0 \leq f_{\text{uf}} < 1$ (Fig. 1C), a perfectly folded and bound protein has no fitness advantage over a folded but unbound protein, and thus this case entails direct selection only for folding. These proteins are harmful to the cell in the misfolded state (e.g., due to aggregation or significant costs of degrading unfolded proteins [4–7]), while binding provides no intrinsic fitness advantage (the protein may have other, functional binding interfaces). Finally, it is also possible to have distinct selection pressures on both binding and folding. This occurs when $0 < f_{\text{ub}} < 1$ and $0 \leq f_{\text{uf}} < 1$ (Fig. 1D).

It is straightforward to generalize our three-state model to proteins with additional structural states (other local minima on the folding energy landscape, other binding modes) and allow for simultaneous adaptation at multiple binding interfaces. Furthermore, the fitness landscape in Eq. 3 can be made an arbitrary nonlinear function of state probabilities. However, these more complex scenarios would still share the essential features of our basic model: coupling between folding and binding traits and sharp fitness thresholds between bound/unbound and folded/unfolded states. Thus our qualitative conclusions do not depend on the specific model in Eq. 3.

Epistasis and local maxima. For protein sequences of length L with an alphabet of size k , each of the k^L possible genotypes is projected onto the two-dimensional trait space of E_f and E_b (Eq. 2) and connected to $L(k - 1)$ immediate mutational neighbors, forming a network of states that the population must traverse (a simple example is shown in Fig. 1E). Adaptive dynamics are determined by the interplay between the structure of the fitness landscape and the distribution of genotypes in trait space.

This interplay gives rise to the possibility of epistasis and multiple local fitness maxima. Our model is non-epistatic in energy space (Eq. 2). When the fitness contours are straight parallel lines, there can be no sign epistasis on the fitness landscape (Fig. 1F). Magnitude epistasis, on the other hand, is widespread due to the nonlinear dependence of fitness on folding and binding energies. Curved fitness contours, which occur near folding or binding thresholds in our model (Fig. 1B–D), can produce sign epistasis in fitness, giving rise to multiple local fitness maxima in the genotype space (Fig. 1E).

Evolutionary dynamics. We assume that a population encoding the protein of interest evolves in the monomorphic limit: $LNu \log N \ll 1$, where L is the number of residues, N is an effective population size, and u is the per-residue probability of mutation per generation [39] (see SI Methods). In this limit, the entire population has the same genotype at any given time, and the rate of substitution from the current genotype to one of its mutational neighbors is given by Eq. 1 in SI Methods. We use the strong-selection limit of the substitution rate (Eq. 2 in SI Methods), in which the effective population size enters only as an overall time scale. In this regime, deleterious mutations never fix and adaptive paths have a finite number of steps, terminating at a global or local fitness maximum. For compact genomic units such as proteins, the monomorphic condition is generally met in multicellular species, although it may be violated in some unicellular eukaryotes and prokaryotes [40]. Sequential fixation of single mutants is also a typical mode of adaptation in directed evolution experiments [3]. For simplicity, we neglect more complex mutational moves such as indels and recombination.

Far from the binding and folding thresholds the fitness landscape becomes flat (Fig. 1A) and the strong-selection assumption may be violated. To establish the limits of validity for our model, we calculate average selection coefficients of accessible substitutions (defined as $s = \mathcal{F}_{\text{final}}/\mathcal{F}_{\text{initial}} - 1$, where $\mathcal{F}_{\text{initial}}$ and $\mathcal{F}_{\text{final}}$ are the initial and final fitness values of a substitution), both throughout the landscape and at the local maxima (Fig. S1). We observe that for typical values of the effective population size $N \in (10^4, 10^7)$ [40, 41], the selection strengths in the model justify our strong-selection approximation for realistic choices of energy parameters.

Quantitative description of adaptation. Although our model is valid for any adaptive process, for concreteness we focus on a specific but widely-applicable scenario. A population begins as perfectly adapted to binding an original target molecule characterized by an energy matrix ϵ_{b_1} with minimum binding energy $E_{b_1}^{\text{min}}$ (defining a fitness landscape \mathcal{F}_1). The population is then subjected to a selection pressure which favors binding a new target, with energy matrix ϵ_{b_2} and minimum binding energy $E_{b_2}^{\text{min}}$ (fitness landscape \mathcal{F}_2). The adaptive paths are first-passage paths leading from the global maximum on \mathcal{F}_1 to a local or global maximum on \mathcal{F}_2 , with fitness increasing monotonically along each path.

Each adaptive path φ with probability $\Pi[\varphi]$ is a sequence of genotypes connecting initial and final states. Using an exact numerical algorithm (SI Methods) [29, 30], we determine the path-length distribution $\rho(\ell)$, which gives the probability of taking an adaptive path with ℓ amino acid substitutions, and the mean adaptation time \bar{t} . We also introduce S_{path} , the entropy of the adaptive paths:

$$S_{\text{path}} = - \sum_{\varphi} \Pi[\varphi] \log \Pi[\varphi]. \quad (4)$$

The path entropy is maximized when evolution is neutral, resulting in all paths of a given length being accessible and equally likely: $S_{\text{path}} = \bar{\ell} \log L(k-1)$ [30], where $\bar{\ell}$ is the average path length.

We also consider the path density $\psi(\sigma)$, which gives the total probability of reaching a

state σ at any point along a path. When σ is a final state (a local fitness maximum on \mathcal{F}_2), the path density is equivalent to the commitment probability. We calculate the entropy S_{com} of the commitment probabilities as

$$S_{\text{com}} = - \sum_{\text{final states } \sigma} \psi(\sigma) \log \psi(\sigma). \quad (5)$$

Direct selection for binding only. We first focus on the $f_{\text{ub}} = f_{\text{uf}} f_{\text{ub}}$ case in Eq. 3. The geometry of the fitness contours is invariant under overall shifts in the binding energy E_b (Fig. 1B); equivalently, the direction (but not the magnitude) of the selection force ($\vec{\nabla} \log \mathcal{F} / |\vec{\nabla} \log \mathcal{F}|$) does not depend on E_b . Thus without loss of generality, we set $E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}}$ in this section. The contours of constant fitness are parallel to the E_f axis when E_f is low, indicating that, as expected, selection acts only on binding when proteins are sufficiently stable.

However, for marginally stable proteins [12, 13, 42], the fitness contours begin to curve downward, indicating apparent, indirect selection for folding, even though selection acts directly only on the binding trait. Thus, adaptation will produce a trait (more stability) that is neutral at the level of the fitness function simply because it is coupled with another trait (binding) that is under selection. Folding stability can therefore be considered an evolutionary spandrel [20]. Proteins may even be intrinsically unstable ($E_f > 0$) and only fold when bound ($E_f + E_b < 0$), which we refer to as binding-mediated stability [43]. In this regime, the fitness contours approach diagonal lines: selection effectively acts to improve both binding and folding equally (Fig. 1B).

An example realization of evolutionary dynamics in the marginally stable regime is shown in Fig. 2A,B (see Fig. S2 for stable and intrinsically unstable examples, and Fig. S3 for averaged distributions of initial, intermediate, and final states). There is typically just one or two fitness maxima; all maxima are usually accessible (Fig. 2C). For stable proteins, the global maximum almost always coincides with the best-binding genotype and is usually as far as a randomly-chosen genotype from the best-folding genotype (Fig. 2D; two random sequences are separated by $1 - 1/k = 0.8$ for $k = 5$). However, as E_f becomes greater, the average distance between the maxima and the best-binding genotype increases, while the average distance between the maxima and the best-folding genotype decreases, until they meet halfway for intrinsically unstable proteins, where effective selection for binding and folding is equally strong (Fig. 2D). In general the maxima lie on or near the Pareto front [44], defined here as the set of genotypes such that either E_f or E_b cannot be decreased further without increasing the other (the global maximum is always on the front, while local maxima may not be) (Fig. 2A, Fig. S2).

As E_f increases, the average distance between initial and final states for adaptation decreases. As a result the average path length (number of substitutions) decreases as well, although the variance of path lengths is relatively constant over all energies (Fig. 2E). The path entropy per substitution $S_{\text{path}}/\bar{\ell}$ also decreases with E_f , reflecting greater constraints on adaptive paths (note that $S_{\text{path}}/\bar{\ell} = \log L(k - 1) \approx 3.2$ for neutral evolution). Finally, $S_{\text{com}} \approx 0.31$ in the marginally stable regime (Fig. 2F). Since the average number of maxima

is ≈ 1.9 in this regime (Fig. 2C), the maximum value of S_{com} is $\log 1.9 \approx 0.64$, indicating that not all maxima are equally accessible.

Direct selection for folding only. In this regime, $f_{\text{ub}} = 1$ and $0 \leq f_{\text{uf}} < 1$ in Eq. 3. Similar to the previous case, the geometry of the fitness contours and thus most landscape properties are now independent of E_f (Fig. 1C); equivalently, normalized selection force $\vec{\nabla} \log \mathcal{F} / |\vec{\nabla} \log \mathcal{F}|$ does not depend on E_f .

When the nonfunctional binding is weak, the fitness contours are parallel to the E_b axis, indicating that selection acts only on folding (Fig. 1C). However, with increasing binding strength the fitness contours curve such that the effective selection force attempts to improve both binding and folding equally. Thus binding emerges as an evolutionary spandrel in this case. The weak-binding regime yields a single fitness maximum due to the lack of sign epistasis; this maximum predominantly coincides with the best-folding genotype (Fig. 3A). However, once the binding interaction becomes stronger, there is an increased likelihood of multiple local maxima, located between the best-folding and best-binding genotypes.

Depending on the abundance of the old and new ligands in the cell and their binding properties, several adaptive scenarios may take place. First, the best-binding strengths $E_{b_1}^{\text{min}}$ and $E_{b_2}^{\text{min}}$ of the old and new targets may be similar in magnitude. If both are weak, initial and final states are likely to be the best-folding genotype or close to it (Fig. 3A); in this case, there is a high probability that no adaptation will occur (Fig. 3B). When $E_{b_1}^{\text{min}}$ and $E_{b_2}^{\text{min}}$ are both low, adaptation usually occurs to accommodate the binding specificity of the new ligand (Fig. 3B, Fig. S4A). Surprisingly, we see that proteins frequently evolve stronger binding at the expense of folding (bottom panel of Fig. S4A). This happens due to the constraints of the genotype-phenotype map: not enough genotypes are available to optimize both traits simultaneously.

It is also possible to gain or lose binding affinity at the nonfunctional interface through adaptation. In the first case, the new target has stronger binding than the old one ($E_{b_2}^{\text{min}} < E_{b_1}^{\text{min}}$). Thus the initial state is the best-folding genotype or close to it, and the protein adapts toward a genotype with intermediate folding and binding (Fig. S4B). As before, adaptation is tightly constrained by the genotype-phenotype map, sacrificing the trait (folding stability) under direct selection in order to affect the spandrel (nonfunctional binding interaction). Effectively, the protein switches from being “self-reliant” to needing a binding partner. In the second case ($E_{b_1}^{\text{min}} < E_{b_2}^{\text{min}}$), the dynamics is opposite: the protein loses its nonfunctional binding interface and becomes self-reliant (Fig. S4C). Thus proteins may acquire or lose binding interfaces depending on the availability of ligands that can participate in binding-mediated stability. If the protein’s stability becomes suboptimal due to an environmental change, its stability may be restored not only through stabilizing mutations, but also by developing a novel binding interface.

Divergent evolutionary fates. In the region where the fitness contours in Fig. 1C are curved, it is possible to have two or more local maxima accessible to adaptation, with at least one having negative E_b (strong binding) and at least one having positive E_b (negligible binding) (see Fig. 3C,D for an example landscape). The selection streamlines are divergent in this regime (Fig. 1C). Thus a protein has two fates available to it: one in which it evolves

to bind the target and another in which it does not. The eventual fate of the protein is determined by random mutation events. Indeed, the distribution of final states is strongly bimodal (Fig. 3E), yielding a sizable probability of divergent fates across a range of binding energies (Fig. 3F).

Simultaneous selection for binding and folding. Finally we consider a general case in which $0 < f_{ub} < 1$ and $0 \leq f_{uf} < 1$ in Eq. 3 (Fig. 1D). The fitness landscape is divided into two regions by a straight diagonal contour with fitness f_{ub} and slope -1 . Below this contour, the landscape is qualitatively similar to the case of selection for binding only (Fig. 1B), while above the contour the landscape resembles that of the folding-only selection scenario (Fig. 1C). Thus evolutionary dynamics for proteins with favorable binding and folding energies will largely resemble the case of selection for binding only. However, a qualitatively different behavior will be observed if the distribution of genotypes straddles the diagonal contour (Fig. 4). This will occur when initial folding stability is marginal and initial binding is unfavorable. In this case, selection streamlines around the diagonal contour (Fig. 1D) and the genotype-phenotype map tightly constrain the adaptive paths to gain extra folding stability first, and then lose it as the binding function is improved.

Tempo and rhythm of adaptation. The strength of selection is the primary determinant of the average adaptation time \bar{t} . If the selection coefficient s is small (but $Ns > 1$), the substitution rate $W(\sigma'|\sigma)$ in SI Methods Eq. 1 is proportional to s . Thus, as selection becomes exponentially weaker for lower energies (Fig. S1), adaptation becomes exponentially slower. The distribution of the total adaptation time over an adaptive path is highly nonuniform. For example, in the case of selection for binding only and a marginally stable protein, the adaptation time is concentrated at the end of the path, one mutation away from the final state (Fig. S5A,B). Substitutions at the beginning of the path occur quickly because there are many possible beneficial substitutions and because selection is strong; in contrast, at the end of the path adaptation slows down dramatically as beneficial mutations are depleted and selection strength weakens. This behavior is observed in most of the other model regimes as well.

The exception to this pattern occurs in the case of selection for both binding and folding in marginally-stable and marginally-bound proteins, due to the unique contour geometry (Fig. 1D). As the adaptive paths wrap around the diagonal contour in the region of high E_b and low E_f , the landscape flattens, making selection weaker and substitutions slower (Fig. S5C). Thus most of the waiting occurs in the middle of the path rather than the end (Fig. S5D). Adaptation accelerates toward the end of the path as the strength of selection increases again. If the intermediate slow-down is significant enough, a protein may not have time to complete the second half of its path before environmental conditions change, so that it will never evolve the new binding function.

Discussion

Protein folding and binding as evolutionary spandrels. In the decades since Gould and Lewontin’s paper [20], the existence of evolutionary spandrels has emerged as a critical

evolutionary concept. There are many possible scenarios in which spandrels can evolve [20, 21], although two key mechanisms are neutral processes, such as genetic drift and biases in mutation and recombination [45], and indirect selection arising from coupled traits. Here we have focused on the latter, which we expect to be more important on short time scales.

It has been previously argued that the marginal stability of most proteins may be an evolutionary spandrel that evolved due to mutation-selection balance [3, 12, 13]. We suggest more broadly that having folding stability at all may be a spandrel for proteins with no misfolding toxicity. Even more striking is the possibility that some binding interactions may be spandrels that evolved solely to stabilize proteins with toxic misfolding; this would significantly affect our interpretation of data on proteome-wide interactions [46]. In particular, we expect more widespread nonfunctional interactions among proteins with less intrinsic stability. Indeed, protein abundance is believed to correlate positively with stability ($-E_f$) to explain the observed negative correlation of abundance with evolutionary rate [18, 19]. Furthermore, models of protein-protein interaction networks imply that protein abundance also correlates negatively with the number of interactions [16]. Together these argue that stability should indeed be negatively correlated with the number of interactions. Experiments on specific proteins also support this finding: for example, destabilizing mutations in *E. coli* dihydrofolate reductase were found to be compensated at high temperature by protein binding, which protected against toxic aggregation [28]. Previously the role of binding-mediated stability has been primarily discussed in the context of intrinsically disordered proteins [47], described by the high E_f regime of our model.

Pareto optimization of proteins. The Pareto front is a useful concept in problems of multi-objective optimization [44]. The Pareto front in our model consists of the protein sequences along the low E_f , low E_b edge of the genotype distribution (see e.g. Fig. 2A). Pareto optimization assumes that all states on the front are valid final states for adaptation; this in turn implies that fitness has linear dependence on the individual traits. However, nonlinear fitness functions with saturation effects will confound this assumption. Our model shows how this nonlinearity leads to a small subset of true final states on or even off the front. Thus Pareto optimization does not capture a key feature of the underlying biophysics, providing only a rough approximation to the true dynamics.

Epistasis and evolutionary predictability. Our results also shed light on the role of epistasis – the correlated effects of mutations at different sites – in protein evolution. Epistasis underlies the ruggedness of fitness landscapes [31, 32]. Magnitude epistasis is widespread in our model, while sign epistasis only arises in regions where the fitness contours are curved (Fig. 1E,F). This picture is qualitatively consistent with studies of empirical fitness landscapes [32] and with directed evolution experiments [3].

Epistasis determines the predictability of evolution, an issue of paramount importance in biology [25, 33, 34]. In most cases considered here, limited sign epistasis gives rise to less predictable intermediate pathways (high S_{path}) but highly predictable final outcomes (low S_{com}).

However, there are two major exceptions to this pattern. First, proteins with a binding interaction under no direct selection may have multiple local maxima, some with strong and

others with weak binding (Fig. 3). Here both intermediate pathways and final states are unpredictable – pure chance, in the form of random mutations, drives the population to one binding fate or the other. The second exception occurs in proteins with direct selection for both binding and folding. Here there is usually a single maximum, but the adaptive paths are tightly constrained in energy space (Fig. 4). Thus evolution of proteins with both functional binding and deleterious misfolding, which should include a large fraction of natural proteins, is highly predictable at the level of energy traits.

Methods

Energetics of protein folding and binding. Folding energetics are probed experimentally and computationally by measuring the changes in E_f resulting from single point mutations. Since these changes are observed to be universally distributed over many proteins [8], we sample entries of ϵ_f from a Gaussian distribution with mean 1.25 kcal/mol and standard deviation 1.6 kcal/mol. For the reference sequence σ_{ref} , $\epsilon_f(i, \sigma_{\text{ref}}^i) = 0$ for all $i \in \{1, \dots, L\}$, such that $E_f(\sigma_{\text{ref}}) = E_f^{\text{ref}}$. The parameter E_b^{min} is defined as the binding energy of the genotype σ_{bb} with the lowest E_b : $\epsilon_b(i, \sigma_{\text{bb}}^i) = 0$ for all $i \in \{1, \dots, L\}$. Since binding hotspot residues typically have a 1–3 kcal/mol penalty for mutations away from the wild-type amino acid [36, 37], we sample the other entries of ϵ_b from an exponential distribution defined in the range of $(1, \infty)$ kcal/mol, with mean 2 kcal/mol. This distribution is consistent with alanine-scanning experiments which probe energetics of amino acids at the binding interface [48]. We consider $L = 6$ hotspot residues and a reduced alphabet of $k = 5$ amino acids (grouped into negative, positive, polar, hydrophobic, and other), resulting in $5^6 = 15625$ possible genotypes. Our population genetics model and the algorithm for exact calculation of adaptive path statistics are available in SI Methods.

Acknowledgments

A.V.M. acknowledges support from National Institutes of Health (R01 HG004708) and an Alfred P. Sloan Research Fellowship.

References

- [1] Creighton TE (1992) *Proteins: Structures and Molecular Properties*. (W.H. Freeman and Company, New York).
- [2] Campbell RE et al. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci USA* 99:7877–7882.
- [3] Bloom JD, Arnold FH (2009) In the light of directed evolution: Pathways of adaptive protein evolution. *Proc Natl Acad Sci USA* 106:9995–10000.

- [4] Bucciantini M et al. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416:507–511.
- [5] Drummond DA, Wilke CO (2008) Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134:341–352.
- [6] Geiler-Samerotte KA et al. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc Natl Acad Sci USA* 108:680–685.
- [7] Bershtein S, Mu W, Serohijos AWR, Zhou J, Shakhnovich EI (2013) Protein quality control acts on folding intermediates to shape the effects of mutations on organismal fitness. *Mol Cell* 49:133–144.
- [8] Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS (2007) The stability effects of protein mutations appear to be universally distributed. *J Mol Biol* 369:1318–1332.
- [9] Tokuriki N, Stricher F, Serrano L, Tawfik DS (2008) How protein stability and new functions trade off. *PLoS Comput Biol* 4:e1000002.
- [10] Wang X, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* 320:85–95.
- [11] Sun SB et al. (2013) Mutational analysis of 48g7 reveals that somatic hypermutation affects both antibody stability and binding affinity. *J Am Chem Soc* 135:9980–9983.
- [12] Taverna DM, Goldstein RA (2002) Why are proteins marginally stable? *Proteins* 46:105–109.
- [13] Zeldovich KB, Chen P, Shakhnovich EI (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci USA* 104:16152–16157.
- [14] Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. *Proc Natl Acad Sci USA* 103:5869–5874.
- [15] Johnson ME, Hummer G (2011) Nonspecific binding limits the number of proteins in a cell and shapes their interaction networks. *Proc Natl Acad Sci USA* 108:603–608.
- [16] Heo M, Maslov S, Shakhnovich EI (2011) Topology of protein interaction network shapes protein abundances and strengths of their function and nonspecific interactions. *Proc Natl Acad Sci USA* 108:4258–4263.
- [17] DePristo MA, Weinreich DM, Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet* 6:678–687.
- [18] Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed proteins evolve slowly. *Proc Natl Acad Sci USA* 102:14338–14343.

- [19] Serohijos AWR, Rimas Z, Shakhnovich EI (2012) Protein biophysics explains why highly abundant proteins evolve slowly. *Cell Rep* 2:249–256.
- [20] Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proc R Soc Lond B* 205:581–598.
- [21] Pigliucci M, Kaplan J (2000) The fall and rise of Dr Pangloss: adaptationism and the Spandrels paper 20 years later. *Trends Ecol Evol* 15:66–77.
- [22] Weiss MA et al. (2002) Protein structure and the spandrels of San Marco: Insulin’s receptor-binding surface is buttressed by an invariant leucine essential for its stability. *Biochemistry* 41:809–819.
- [23] Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* 12:767–780.
- [24] Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. *Nat Rev Genet* 11:572–582.
- [25] Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312:111–114.
- [26] Chou HH, Chiu HC, Delaney NF, Segrè D, Marx CJ (2011) Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332:1190–1192.
- [27] Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ (1999) Different trajectories of parallel evolution during viral adaptation. *Science* 285:422–424.
- [28] Bershtein S, Mu W, Shakhnovich EI (2012) Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations. *Proc Natl Acad Sci USA* 109:4857–4862.
- [29] Manhart M, Morozov AV (2013) Path-based approach to random walks on networks characterizes how proteins evolve new functions. *Phys Rev Lett* 111:088102.
- [30] Manhart M, Morozov AV (2014) in *First-Passage Phenomena and Their Applications*, eds. Metzler R, Oshanin G, Redner S. (World Scientific, Singapore).
- [31] Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445:383–386.
- [32] Szendro IG, Schenk MF, Franke J, Krug J, de Visser JA (2013) Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech.* p. P01005.
- [33] Gould SJ (1990) *Wonderful Life: The Burgess Shale and the Nature of History*. (W. W. Norton and Company, New York, USA).
- [34] Lobkovsky AE, Koonin EV (2012) Replaying the tape of life: quantification of the predictability of evolution. *Front Gene* 3:246.

- [35] Kimura M (1983) *The Neutral Theory of Molecular Evolution*. (Cambridge University Press, Cambridge, UK).
- [36] Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* 267:383–386.
- [37] Moreira IS, Fernandes PA, Ramos MJ (2007) Hot spots — a review of the protein-protein interface determinant amino-acid residues. *Proteins* 68:803–812.
- [38] Wells JA (1990) Additivity of mutational effects in proteins. *Biochemistry* 29:8509–8517.
- [39] Champagnat N (2006) A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stoch Proc Appl* 116:1127–1160.
- [40] Lynch M (2007) *The Origins of Genome Architecture*. (Sinauer, Sunderland).
- [41] Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 10:195–205.
- [42] Kumar MD et al. (2006) ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res* 34:D204–D206.
- [43] Dixit PD, Maslov S (2013) Evolutionary capacitance and control of protein stability in protein-protein interaction networks. *PLoS Comput Biol* 9:e1003023.
- [44] Shoval O et al. (2012) Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* 336:1157–1160.
- [45] Lynch M (2007) The evolution of genetic networks by non-adaptive processes. *Nat Rev Genet* 8:803–813.
- [46] Stark C et al. (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 34(Database issue):D535–D539.
- [47] Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19:31–38.
- [48] Thorn KS, Bogan AA (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* 17:284–285.
- [49] Kimura M (1962) On the probability of fixation of mutant genes in a population. *Genetics* 47:713–719.
- [50] Istomin AY, Gromiha MM, Vorov OK, Jacobs DJ, Livesay DR (2008) New insight into long-range nonadditivity within protein double-mutant cycles. *Proteins* 70:915–924.

Figures

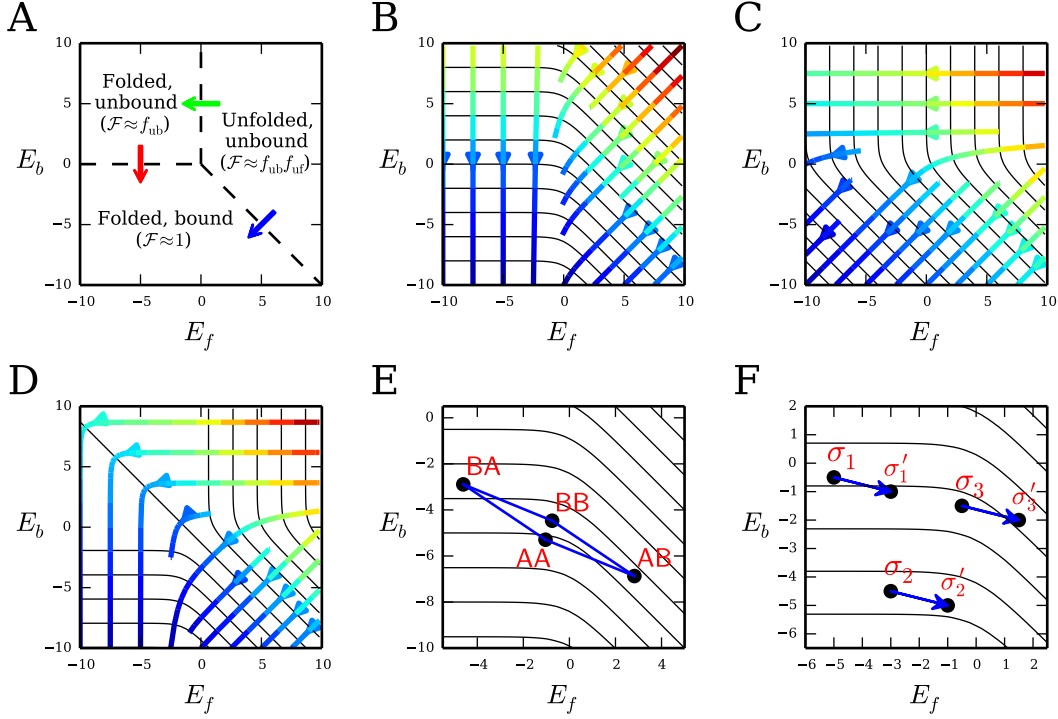


Figure 1: **Fitness, selection, and epistasis in energy trait space.** (A) Phase diagram of protein structural states. Dashed lines separate structural phases of the protein corresponding to plateaus on the fitness landscape; arrows represent the folding transition (green), binding transition (red), and the coupled folding-binding transition (blue). Fitness landscapes $\mathcal{F}(E_f, E_b)$ with direct selection (B) for binding only ($f_{ub} = f_{uf} = 0$), (C) for folding only ($f_{ub} = 1, f_{uf} = 0$), and (D) for both binding and folding ($f_{ub} = 0.9, f_{uf} = 0$). Black contours indicate constant fitness values. The contours are uniformly spaced in energy space; fitness differences between adjacent contours are not all equal. Streamlines indicate the direction of the selection “force” $\vec{\nabla} \log \mathcal{F}$, with color showing its magnitude (decreasing from red to blue). (E) Projection of a genotype distribution and mutational network into energy space for $L = 2$ and a two-letter ($k = 2$) alphabet. (F) Blue arrows indicate the same mutation on different genetic backgrounds. When the fitness contours are straight, the mutation is beneficial regardless of the background (σ_1 or σ_2). However, with curved contours, the same mutation can become deleterious ($\sigma_3 \rightarrow \sigma'_3$), indicative of sign epistasis. Sign epistasis from curved contours can give rise to multiple local fitness maxima (e.g., AA and BB in (E)).

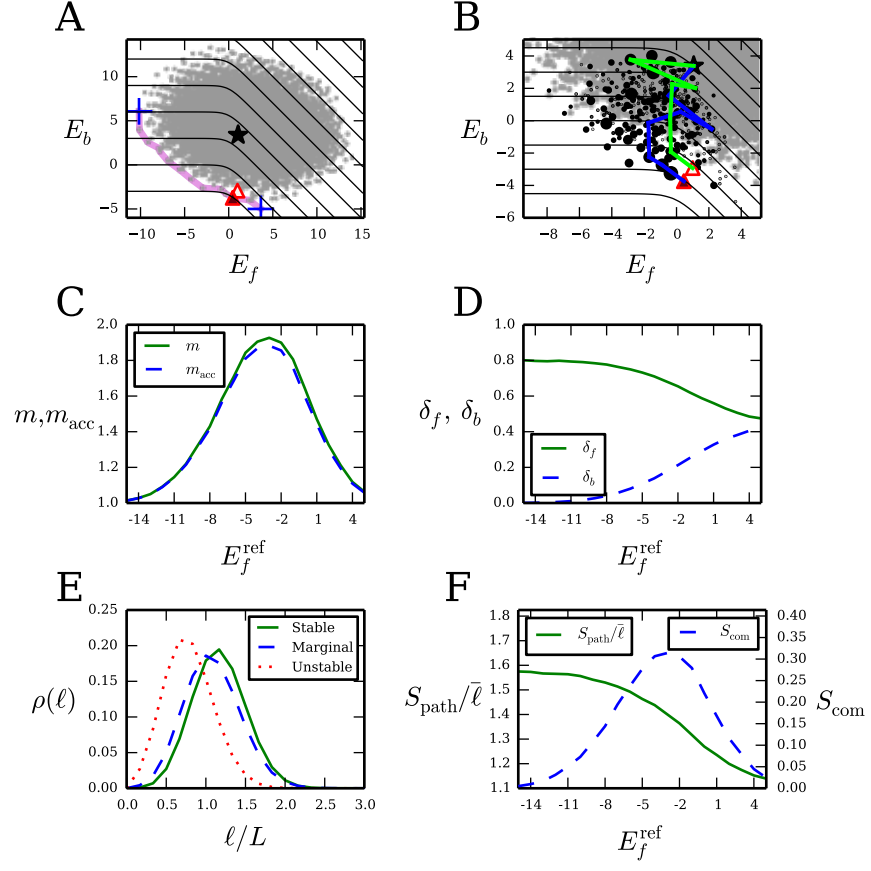


Figure 2: **Properties of adaptation with direct selection for binding only.**

(A) Global distribution of folding and binding energies for all $k^L = 5^6$ genotypes in a single realization of the model with a marginally stable protein ($E_f^{\text{ref}} = -3$ kcal/mol). The black star indicates the initial state for adaptation (global maximum on \mathcal{F}_1), red triangles indicate local fitness maxima on \mathcal{F}_2 , shaded according to their commitment probabilities $\psi(\sigma)$, and the blue crosses indicate best-folding and best-binding genotypes. The magenta line connects genotypes on the Pareto front, and the black contours indicate constant fitness \mathcal{F}_2 . (B) The region of energy space accessible to adaptive paths, zoomed in from (A). Example paths are shown in blue and green; black circles indicate intermediate states along paths, sized proportional to their path density $\psi(\sigma)$; small gray circles are genotypes inaccessible to adaptation. (C) Average number m of local fitness maxima (solid, green) and average number m_{acc} of local maxima accessible to adaptation (dashed, blue) versus E_f^{ref} . The average number of maxima is greatest at $E_f^{\text{ref}} \approx -3$ kcal/mol, where multiple local maxima are separated by ≈ 2.23 substitutions on average. (D) Average per-residue Hamming distance between the maxima and the best-folding genotype (δ_f ; solid, green) and the best-binding genotype (δ_b ; dashed, blue) versus E_f^{ref} . (E) Average distributions $\rho(\ell)$ of path lengths (number of substitutions) ℓ for stable proteins ($E_f^{\text{ref}} = -15$ kcal/mol), marginally stable proteins ($E_f^{\text{ref}} = -3$ kcal/mol), and intrinsically unstable proteins ($E_f^{\text{ref}} = 5$ kcal/mol). (F) Per-substitution path entropy $S_{\text{path}}/\bar{\ell}$ (solid, green) and entropy of commitment probabilities S_{com} (dashed, blue) versus E_f^{ref} . Panel (E) is averaged over 10^5 realizations of the model; all other averages are taken over 10^4 realizations. In all panels $f_{\text{ub}} = f_{\text{uf}} = 0$ and $E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -5$ kcal/mol.

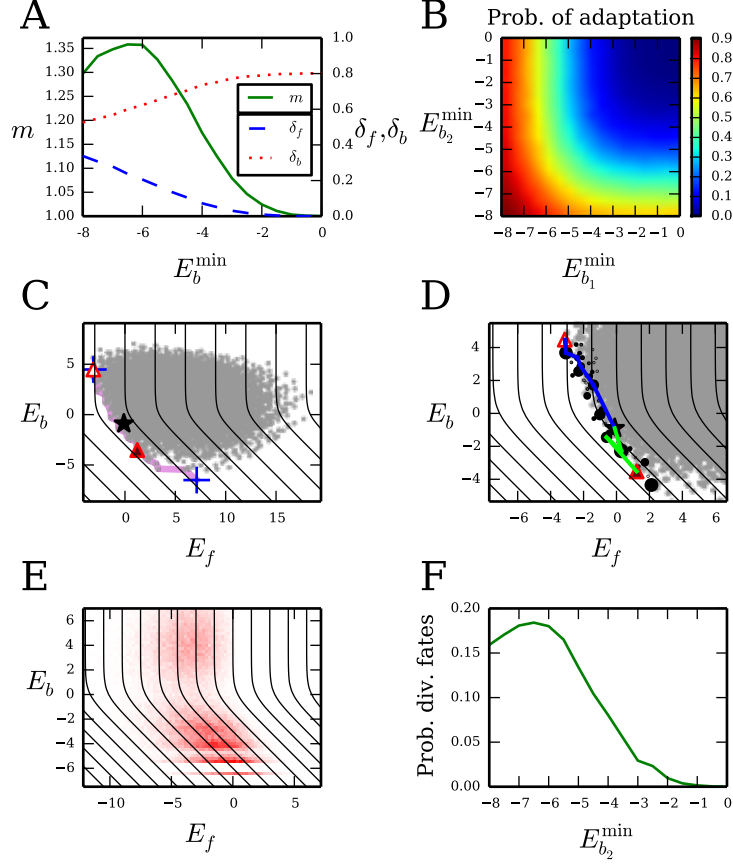


Figure 3: **Properties of adaptation with direct selection for folding only.** (A) The average number of local maxima m (solid, green) and their average per-residue Hamming distance from the best-folding (δ_f ; dashed, blue) and the best-binding (δ_b ; dotted, red) genotypes versus E_b^{\min} . (B) Probability that adaptation occurs when the binding target is changed (i.e., the initial state is not coincident with any of the final states), as a function of $E_{b_1}^{\min}$ and $E_{b_2}^{\min}$. (C,D) Example landscape with divergent binding fates: there are two accessible local maxima, one with $E_b < 0$ (favorable binding, $\psi(\sigma) = 0.6$) and the other with $E_b > 0$ (negligible binding, $\psi(\sigma) = 0.4$). All symbols are the same as in Fig. 2A,B. (E) Average distribution of local maxima, weighted by their commitment probabilities. The average commitment entropy for realizations with divergent fates is $S_{\text{com}} \approx 0.43$. In (C)–(E) we used $E_{b_1}^{\min} = E_{b_2}^{\min} = -6.5$ kcal/mol. (F) The probability of having divergent fates versus $E_{b_2}^{\min} = E_{b_1}^{\min}$. Panel (E) is averaged over 10^5 realizations of the model; all other averages are taken over 10^4 realizations. In all panels $f_{\text{ub}} = 1$, $f_{\text{uf}} = 0$, and $E_f^{\text{ref}} = 0$ kcal/mol.

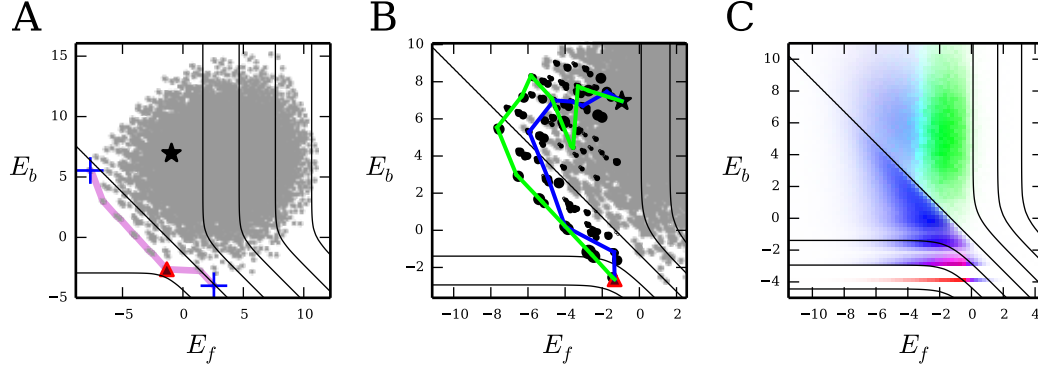


Figure 4: **Properties of adaptation with direct selection for both folding and binding.** (A, B) Distribution of folding and binding energies in an example landscape for a marginally stable and marginally bound protein; all symbols are the same as in Fig. 2A,B. (C) Landscape averaged over 10^5 realizations. Distribution of initial states is shown in green, intermediate states in blue (weighted by their path densities), and final states in red (weighted by their commitment probabilities). In all panels $f_{\text{ub}} = 0.9$, $f_{\text{uf}} = 0$, and $E_f^{\text{ref}} = E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -4$ kcal/mol.

Supplementary Material:

Protein folding and binding can emerge as evolutionary spandrels through structural coupling

Michael Manhart¹ and Alexandre V. Morozov^{1,2}

¹*Department of Physics and Astronomy and* ²*BioMaPS Institute for Quantitative Biology,*
Rutgers University, Piscataway, New Jersey 08854, USA

Supplementary Methods

Population genetics model. In the monomorphic limit, the population is described by a single point in genotype space [39]. The population evolves over time via mutations that arise sequentially and either fix or disappear. Each fixation event leads to an amino acid substitution in the entire population. The rate of making a substitution from genotype σ to genotype σ' is given by [35]

$$W(\sigma'|\sigma) = Nu \phi(\sigma'|\sigma), \quad (\text{S1})$$

where N is the effective population size, u is the mutation rate, and $\phi(\sigma'|\sigma)$ is the probability of a single σ' mutant fixing in a population of wild-type σ . Typically the fixation probability depends only on the relative selection coefficient $s = \mathcal{F}(\sigma')/\mathcal{F}(\sigma) - 1$ between the two genotypes, where $\mathcal{F}(\sigma)$ is the fitness of genotype σ . For example, in the Wright-Fisher model, $\phi(s) = (1 - e^{-2s})/(1 - e^{-2Ns})$, where N is the effective population size [49]. In the strong-selection limit ($N|s| \gg 1$),

$$\phi(s) \approx \begin{cases} 1 - e^{-2s} & \text{for } s > 0 \\ 0 & \text{for } s < 0 \end{cases} \quad (\text{S2})$$

Thus the effective population size N sets the overall time scale $(Nu)^{-1}$ of substitutions but does not affect fixation probabilities.

Statistics of adaptive paths. We calculate statistical properties of the adaptive paths using a transfer matrix-like algorithm [29, 30]. Let \mathcal{S} be the set of all genotypes accessible to adaptation, and let \mathcal{S}_f be the set of final state genotypes (e.g., local fitness maxima). Define $W(\sigma'|\sigma)$ as the rate of making a substitution from genotype σ to genotype σ' (e.g., given by Eq. S1). The rate matrix defines $\theta(\sigma) = (\sum_{\text{nn } \sigma' \text{ of } \sigma} W(\sigma'|\sigma))^{-1}$, the mean waiting time in genotype σ before a substitution occurs, where the sum is over all genotypes σ' one mutation away from σ (nearest mutational neighbors, “nn”). The substitution rates also determine the probability $Q(\sigma'|\sigma) = W(\sigma'|\sigma)\theta(\sigma)$ of making the substitution $\sigma \rightarrow \sigma'$, given that a substitution occurs out of σ .

For each substitution ℓ and intermediate genotype σ , we calculate $P_\ell(\sigma)$, the total probability of all paths that end at σ in ℓ substitutions; $T_\ell(\sigma)$, the total average time of all such paths; and $\Gamma_\ell(\sigma)$, their total entropy. These quantities obey the following recursion relations:

$$\begin{aligned} P_\ell(\sigma') &= \sum_{\text{nn } \sigma \text{ of } \sigma'} Q(\sigma'|\sigma) P_{\ell-1}(\sigma), \\ T_\ell(\sigma') &= \sum_{\text{nn } \sigma \text{ of } \sigma'} Q(\sigma'|\sigma) [T_{\ell-1}(\sigma) + \theta(\sigma) P_{\ell-1}(\sigma)], \\ \Gamma_\ell(\sigma') &= \sum_{\text{nn } \sigma \text{ of } \sigma'} Q(\sigma'|\sigma) [\Gamma_{\ell-1}(\sigma) - (\log Q(\sigma'|\sigma)) P_{\ell-1}(\sigma)], \end{aligned} \quad (\text{S3})$$

where $P_0(\sigma) = 1$ if σ is the initial state and $P_0(\sigma) = 0$ otherwise, and $T_0(\sigma) = \Gamma_0(\sigma) = 0$ for all $\sigma \in \mathcal{S}$. The final states $\sigma \in \mathcal{S}_f$ are treated as absorbing to ensure that only first-passage paths are counted. We use these transfer matrix objects to calculate the path ensemble quantities described in the text:

$$\begin{aligned} \rho(\ell) &= \sum_{\sigma \in \mathcal{S}_f} P_\ell(\sigma), & \psi(\sigma) &= \sum_{\ell=1}^{\Lambda} P_\ell(\sigma), \\ \bar{t} &= \sum_{\ell=1}^{\Lambda} \sum_{\sigma \in \mathcal{S}_f} T_\ell(\sigma) = \sum_{\ell=1}^{\Lambda} \tau(\ell) = \sum_{\sigma \in \mathcal{S}} \tau(\sigma), & \tau(\sigma) &= \sum_{\ell=1}^{\Lambda} \theta(\sigma) P_\ell(\sigma), \\ S_{\text{path}} &= \sum_{\ell=1}^{\Lambda} \sum_{\sigma \in \mathcal{S}_f} \Gamma_\ell(\sigma), & \tau(\ell) &= \sum_{\sigma \in \mathcal{S}} \theta(\sigma) P_\ell(\sigma). \end{aligned} \quad (\text{S4})$$

The sums are calculated up to a path length cutoff Λ , which we choose such that $1 - \sum_{\ell=1}^{\Lambda} \rho(\ell) < 10^{-6}$. Note that the calculations for the state-dependent quantities $\psi(\sigma)$ and $\tau(\sigma)$ are simplified in this model (compared to more general cases [29, 30]) since the strong-selection dynamics prevents the population from traversing loops in genotype space. The time complexity of the algorithm scales as $\mathcal{O}(\gamma N \Lambda)$ [29], where γ is the average connectivity and N is the total size of the state space. For genotypic sequences of length L and an alphabet of size k , $\gamma \sim L(k-1)$ and $N \sim k^L$.

Validity of the additive energy model. Double mutant experiments indicate that the additive energy model is a good approximation for residues that are not in direct physical contact [38, 50]. For spatially-close residues, the mutational effects are largely “sub-additive” (diminishing-returns magnitude epistasis): two (de)stabilizing mutations combined will still usually be (de)stabilizing, but less so than the sum of their individual effects [38, 50]. For example, Istomin et al. [50] find that while residues separated by more than 6 Å are nearly additive (correlation $R^2 = 0.97$ with a slope of 0.88 between the sum of $\Delta\Delta G$ ’s for two single mutants and $\Delta\Delta G$ for the double mutant), spatially-close residues are substantially sub-additive ($R^2 = 0.84$, slope of 0.54). Nonetheless, in regions with straight contours which

represent most of our fitness landscapes, sub-additive energies cannot produce sign epistasis; substantial deviations from energy sub-additivity are required to create additional local maxima or place significant constraints on adaptive paths. Thus it appears that deviations from energy additivity will not lead to qualitative changes in our model's predictions.

Supplementary Figures

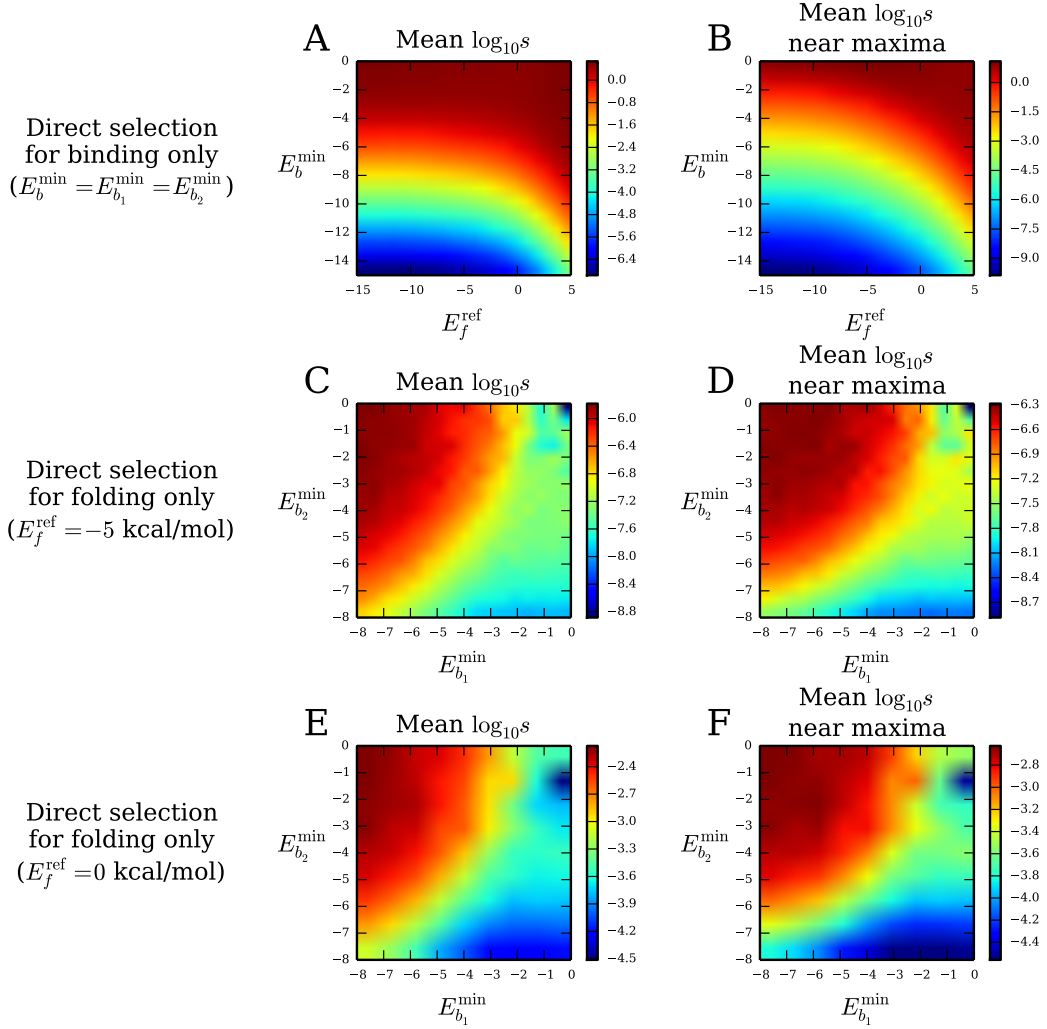


Figure S1: **Average selection strength.** (A) Average $\log_{10} s$ (s is the selection coefficient) of all accessible beneficial substitutions as a function of E_f^{ref} and $E_b^{\min} = E_{b_1}^{\min} = E_{b_2}^{\min}$ in the case of direct selection for binding only ($f_{\text{uf}} = f_{\text{ub}} = 0$). Due to the E_b symmetry of this case (Fig. 1B), we can neglect differences in $E_{b_1}^{\min}$ and $E_{b_2}^{\min}$ without loss of generality. (B) Same as (A) but limited to accessible substitutions that end at local fitness maxima. (C) Average $\log_{10} s$ of all accessible beneficial substitutions as a function of $E_{b_1}^{\min}$ and $E_{b_2}^{\min}$ in the case of selection for folding only ($f_{\text{uf}} = 0$, $f_{\text{ub}} = 1$, $E_f^{\text{ref}} = -5$ kcal/mol). (D) Same as (C) but limited to accessible substitutions that end at local fitness maxima. (E, F) Same as (C, D) but for $E_f^{\text{ref}} = 0$ kcal/mol. Simultaneous selection for both binding and folding yields qualitatively similar results. All data points are averages over 10^4 landscape realizations.

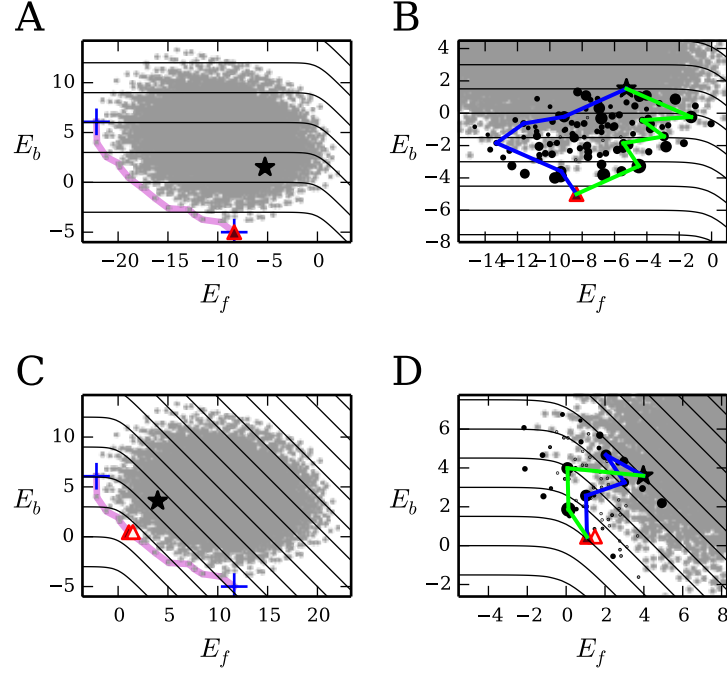


Figure S2: **Example landscapes for stable and intrinsically unstable proteins with direct selection for binding only.** Symbols and randomly generated energy matrices (ϵ_f , ϵ_{b_1} , and ϵ_{b_2}) are the same as in Fig. 2A,B. (A, B) Stable protein ($E_f^{\text{ref}} = -15$ kcal/mol). (C, D) Intrinsically unstable protein ($E_f^{\text{ref}} = 5$ kcal/mol). As in Fig. 2A,B, $f_{\text{ub}} = f_{\text{uf}} = 0$ and $E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -5$ kcal/mol.

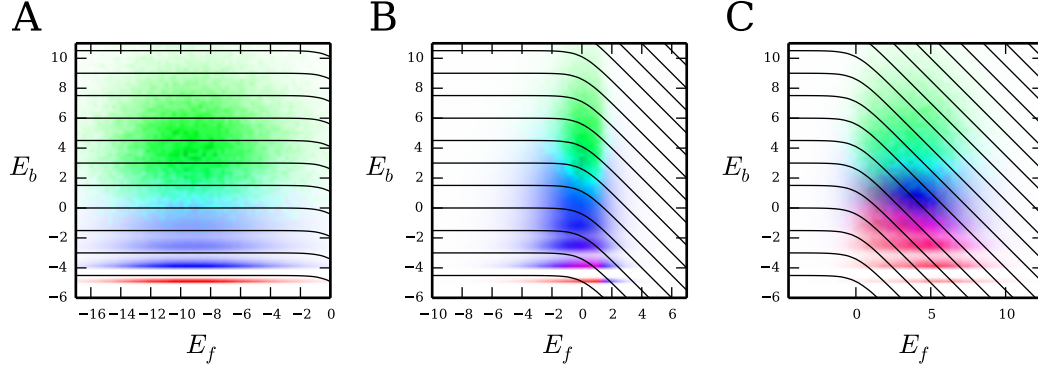


Figure S3: **Average landscapes for direct selection for binding only.** As in Fig. 4C, the distribution of initial states is shown in green, intermediate states in blue (weighted by their path densities), and final states in red (weighted by their commitment probabilities). (A) Stable proteins ($E_f^{\text{ref}} = -15$ kcal/mol). (B) Marginally stable proteins ($E_f^{\text{ref}} = -3$ kcal/mol). (C) Intrinsically unstable proteins ($E_f^{\text{ref}} = 5$ kcal/mol). All landscapes are averaged over 10^5 realizations. As in Fig. 2A,B, $f_{\text{ub}} = f_{\text{uf}} = 0$ and $E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -5$ kcal/mol.

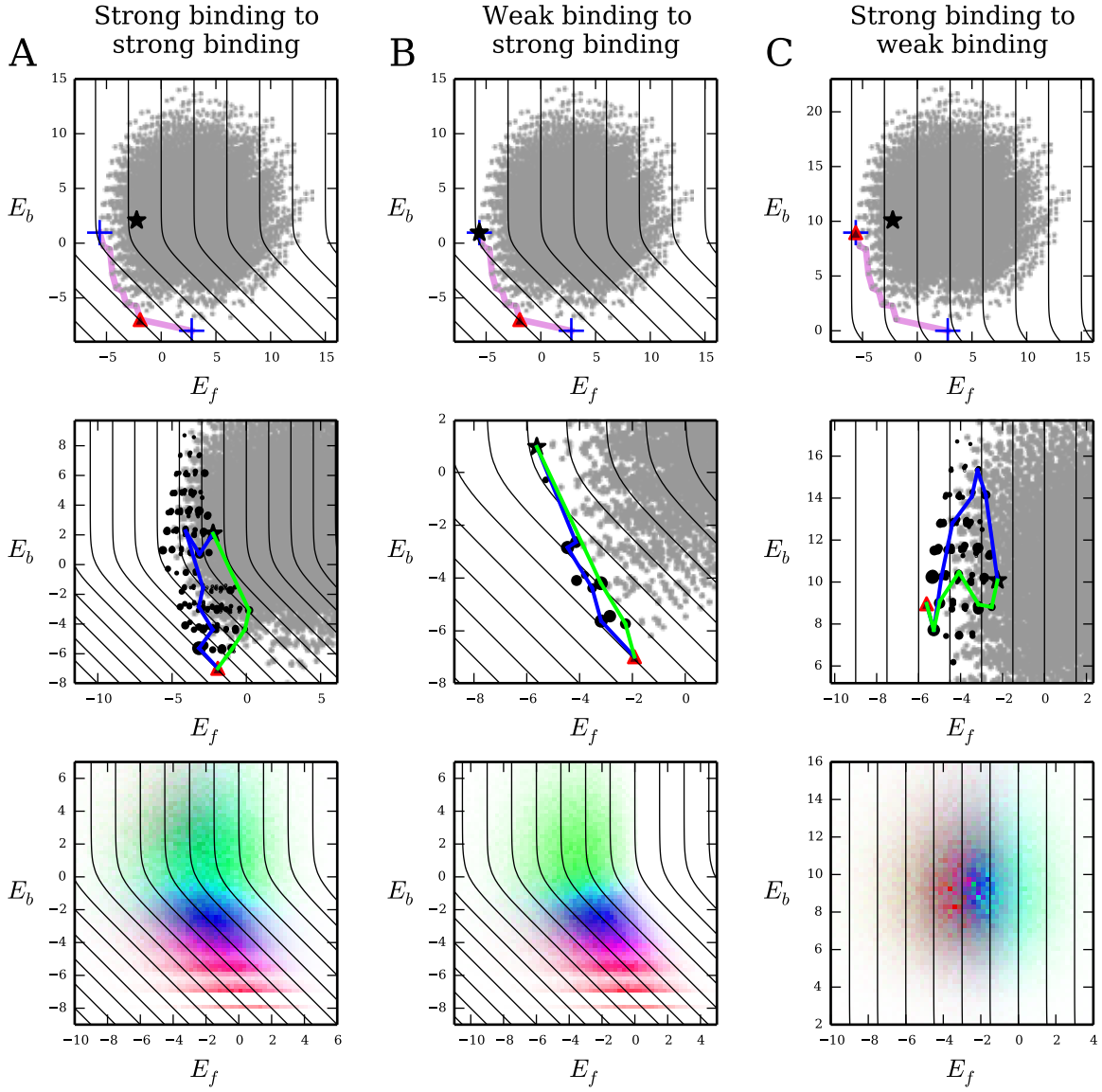


Figure S4: **Example and average landscapes for direct selection for folding only.** Symbols in top and middle panels are the same as in Fig. 2A,B, and the color scheme in the bottom panels is the same as in Fig. 4C and Fig. S3. (A) Strong binding to both old and new targets ($E_{b_1}^{\min} = E_{b_2}^{\min} = -8$ kcal/mol). (B) Weak binding to old target and strong binding to new target ($E_{b_1}^{\min} = 0$ kcal/mol, $E_{b_2}^{\min} = -8$ kcal/mol). (C) Strong binding to old target and weak binding to new target ($E_{b_1}^{\min} = -8$ kcal/mol, $E_{b_2}^{\min} = 0$ kcal/mol). We use $f_{ub} = 1$, $f_{uf} = 0$, and $E_f^{\text{ref}} = 0$ kcal/mol in all cases. In the bottom panels, the landscapes are averaged over 10^5 realizations.

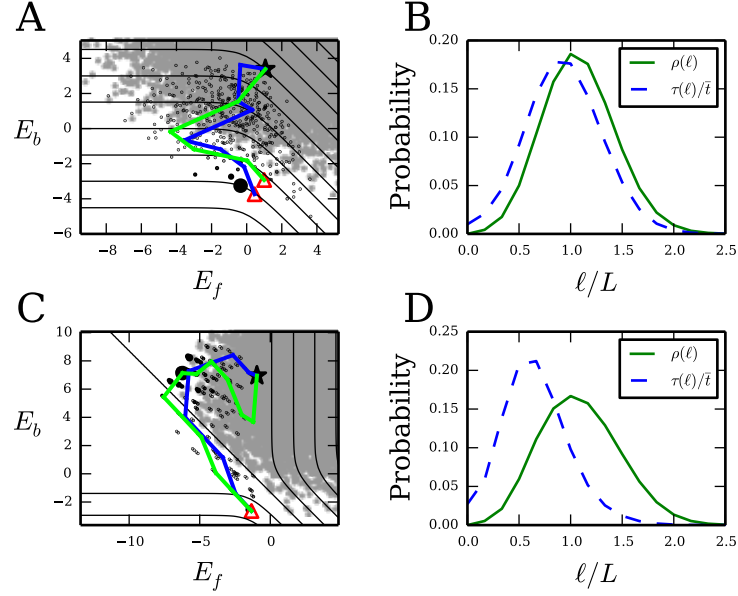


Figure S5: **Distribution of adaptation times over intermediate states.** (A) The same landscape realization as in Fig. 2A,B (selection for binding only on a marginally stable protein), but with each intermediate state σ sized proportional to $\tau(\sigma)$, the average time spent in that state. (B) The probability $\rho(\ell)$ (solid, green) of taking an adaptive path of exactly ℓ substitutions and the average time $\tau(\ell)$ (dashed, blue) spent by paths at the ℓ th substitution, averaged over 10^5 realizations with $f_{\text{ub}} = f_{\text{uf}} = 0$, $E_f^{\text{ref}} = -3$ kcal/mol, and $E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -5$ kcal/mol. (C, D) Same as (A, B), but with the landscape realization used in Fig. 4A,B (selection for both binding and folding, $f_{\text{ub}} = 0.9$, $f_{\text{uf}} = 0$, $E_f^{\text{ref}} = E_{b_1}^{\text{min}} = E_{b_2}^{\text{min}} = -4$ kcal/mol).